J Gastroenterol (2022) 57:1–9
https://doi.org/10.1007/s00535-021-01835-7

The Japanese Society of Gastroenterology

ORIGINAL ARTICLE—ALIMENTARY TRACT

Effects of a novel endoscopic reporting system with voice recognition on the endoscopic procedure time and report preparation time: propensity score matching analysis

Yuta Yokota[1] · Taro Iwatsubo[2] · Toshihisa Takeuchi[3] · Akitoshi Hakoda[1,2] · Yasuki Nakagawa[1] · Kazumi Kawabata[1] · Yuma Inoue[1] · Hayato Miyamoto[1] · Koichi Ikeo[1] · Yuichi Kojima[3] · Junichi Miyazaki[1] · Takashi Abe[1] · Kazuhide Higuchi[2]

Received: 30 August 2021 / Accepted: 20 October 2021 / Published online: 25 October 2021

Abstract
Background With the increase in endoscopic procedures, endoscopists are spending more time creating reports. Although medical reports have largely become electronic, most of the current reporting systems require manual operation. This study aimed to evaluate the efficacy of a novel endoscopic reporting system that uses voice recognition (VR) technology.
Methods We retrospectively reviewed consecutive patients who underwent esophagogastroduodenoscopy between September 2019 and March 2020 at a general hospital in Japan. The novel reporting system, used during endoscopic procedures, is equipped with VR and provides automatic responses by playing back recognized words. Differences in total time spent on the endoscopic procedure and report preparation between the manual entry (ME) and VR groups were evaluated using a propensity score matching method.
Results We enrolled 356 patients: 226 and 130 patients in the ME and VR groups, respectively. Propensity score matching created 101 matched pairs. After matching, the median report preparation time (311 vs. 383 s, $P = 0.009$) and median total time (765 vs. 842 s, $P = 0.053$) in the VR group were shorter than those in the ME group. The VR system independently shortened the total and report preparation times by 156 s (95% confidence interval, − 274 to − 37 s; $P = 0.009$) and 118 s (95% confidence interval, − 220 to − 15 s; $P = 0.023$), respectively, on multiple linear regression analysis.
Conclusions The VR system could save the report preparation time and the total time. This novel system may improve the efficiency of endoscopy-related tasks.

Keywords Voice recognition · Speech recognition · Esophagogastroduodenoscopy · Endoscopic report

Introduction

In recent years, the number of gastrointestinal endoscopies performed worldwide has increased. The fundamental elements of any endoscopic procedure, including findings, therapy, clinical recommendations, and adverse events, should be shared in endoscopy reports. Thus, after the procedure, endoscopists have to spend a substantial amount of time to produce accurate and comprehensive reports. Even more time is spent on report preparation when several lesions are detected or complex treatments are performed. In such cases, problems of incomplete records or lower work efficiency can occur [1].

Endoscopic reporting systems should be electronic to restrict the use of free-text entry and to be based mainly on structured data entry [1]. Structured electronic reports based on terminology are advantageous in that they make it possible to reduce the occurrence of incomplete records, search any database created, and perform statistical analyses [2, 3]. Multiple electronic and structured endoscopic reporting systems have been developed; however, with these systems, endoscopists cannot input their findings

Yuta Yokota and Taro Iwatsubo are co-first authors.

✉ Toshihisa Takeuchi
toshihisa.takeuchi@ompu.ac.jp

[1] Department of Gastroenterology, Hanwa Sumiyoshi General Hospital, Osaka, Japan

[2] 2nd Department of Internal Medicine, Osaka Medical and Pharmaceutical University, Osaka, Japan

[3] Endoscopic Center, Osaka Medical and Pharmaceutical University Hospital, Osaka, Japan

during the endoscopic procedure. Instead, this can only be performed after the procedure [4]. Report preparation after the procedure entails creating extra workloads for endoscopists, which might result in lower work efficiency. Streamlining the input of information on endoscopic findings is warranted to overcome this issue.

Therefore, we developed a novel reporting system for routine esophagogastroduodenoscopy (EGD) that uses voice recognition (VR) technology, which is available during endoscopic procedures. In other medical fields, recording time using VR technology has been reported to be faster than typing [5, 6]. However, it is less known whether an endoscopic reporting system with VR has a higher time efficacy than that of manual entry (ME). Hence, we performed a retrospective observational study to evaluate the efficacy of an endoscopic reporting system that uses VR technology.

Methods

Study design and participants

We enrolled a cohort of consecutive patients who underwent routine EGD at Hanwa Sumiyoshi General Hospital, Osaka, Japan, between September 2019 and March 2020. The total time spent on the endoscopic procedure and report preparation using ME and the VR system was retrospectively evaluated. The exclusion criteria were as follows: examinations performed using endoscopy systems from companies not supporting a VR system, examinations performed by endoscopists who had never used a VR system, examinations of patients with a history of surgical treatment of the upper gastrointestinal tract, an unknown examination time, and therapeutic endoscopy. Moreover, endoscopic procedures or a report preparation that took more than 60 min were excluded, considering the complex examination or the possibility of incomplete registration in the information management system at the time when the procedure or report was completed.

Patient data were retrieved from medical records and an electronic endoscopic database. Informed consent was obtained in the form of an opt-out system on the website. This guaranteed the opportunity for refusal to provide medical information to the study. The Institutional Review Board of Hanwa Sumiyoshi General Hospital approved the study protocol (no. 2020-6, approved on June 5, 2020), and the study was performed in accordance with the Declaration of Helsinki.

Endoscopic procedure

All endoscopists (A, B, C, D, E, F, and G) had performed at least 3000 EGDs, with at least 5 years of experience in performing endoscopic procedures. We administered the following medications 10 min before the examination: ingestion of a mixture consisting of a mucolytic agent (20,000 U pronase, Pronase MS; Kaken Pharmaceutical, Tokyo, Japan), a defoaming agent (80 mg dimethylpolysiloxane syrup, BAROS Antifoaming Oral Solution 2%; Horii Pharmaceutical Ind., Osaka, Japan), and 1 g sodium bicarbonate diluted in 100 ml of tap water for gastric preparation and lidocaine spray (8% Xylocaine Pump Spray; AstraZeneca K.K., Osaka, Japan) containing a local anesthetic for pharyngeal preparation. Midazolam (1.0–5.0 mg) was administered as a sedative when requested by the patients. Additionally, pethidine hydrochloride was administered for sedation if the patient had not been under appropriate sedation with midazolam in the previous examination. Principally, antispasmodic agents were not used. Transoral endoscopy (EG-L600ZW7, Fujifilm Co., Tokyo, Japan) with the LASEREO endoscopic system (Fujifilm Co., Tokyo, Japan) was used for all patients. Image-enhanced endoscopy (blue laser imaging) was performed in esophageal observation with either insertion or withdrawal. Where necessary, chromoendoscopy (including indigo carmine and/or Lugol staining) and a biopsy of lesions suspected of neoplasms were performed. None of the study patients underwent magnifying endoscopy. All endoscopists in this study usually captured 30–40 images, including the esophagus, stomach, and duodenum, during screening EGD. When the abnormal or suspected findings were found, additional images were captured as necessary. The digital images were automatically saved in the endoscopic filing system.

Report preparation

We used an electronic endoscopic information management system (Nexus; Fujifilm Co., Tokyo, Japan) for preparing all reports in Japanese. This system is compatible with the Japan Endoscopy Database (JED) project, which has been promoted with the aim of unifying terminology and establishing a national database for gastroenterological endoscopy [7]. All endoscopists had tried the VR system, were familiar with its usage before the study period, and never knew the purpose of the study during the study period. If both systems connected and unconnected with VR are available, the use of the VR system was the discretion of the endoscopist. In the VR group, the words vocalized by endoscopists during the procedure were temporarily recorded by a VR system and were manually registered as final reports after the procedure (see "VR

system"). By contrast, all reports in the ME group were manually created and registered after the procedure. All report preparations were based on structured data entry with stratification, such as by organ, diagnosis, characteristics, disease classification, and procedure. Free text entry was performed only after the procedure, as necessary, and was restricted to a minimum description of crucial information that could not be expressed in a structured report. In both groups, all manual operations were based on structured data using pull-down menus, checklists, and free-text typing. The registration of the endoscopic report was completed after entering the patient's status and administered premedications and attaching endoscopic pictures and pathology request forms, in addition to the report preparation described above.

VR system

In the VR group, endoscopists used a new endoscopic reporting system with VR technology (voice capture; Rasis Software Service CO., Osaka, Japan) to record endoscopic findings while performing the procedure. This system consists of a data management server, an input terminal (a tablet), and a voice terminal (a headphone with a microphone), as shown in Fig. 1. Adopted terms are compatible with terminology in NEXUS, that is, the JED terminology [8]. The input summary is as follows: first, endoscopists vocalize "start" as soon as the endoscope is inserted to start the procedure. Then the endoscopist vocalizes the endoscopic findings to the voice terminal during the procedure. The VR system automatically recognizes the endoscopist's statement and, at the same time, responds by playing it back (Supplementary Material 1). If the system does not react to the endoscopist's statement, the endoscopist repeats the statement. When the playback from the system is correct, the endoscopist confirms the finding by vocalizing "register." When it is wrong, the incorrect finding is deleted by vocalizing "return," and the endoscopist attempts to repeat entering the finding. Additionally, the VR system can be turned off and on by vocalizing "voice recognition off" and "voice recognition on," respectively, as necessary. Finally, the endoscopist vocalizes "termination" as soon as the endoscope is withdrawn to terminate the procedure. Immediately after termination of the procedure, the registered record is transferred, as a temporary report, to a data administration server. Afterward, the temporary report is checked on the data administration server, and the document is manually modified and added, as necessary. The registration of the endoscopic reports is eventually completed after careful review.

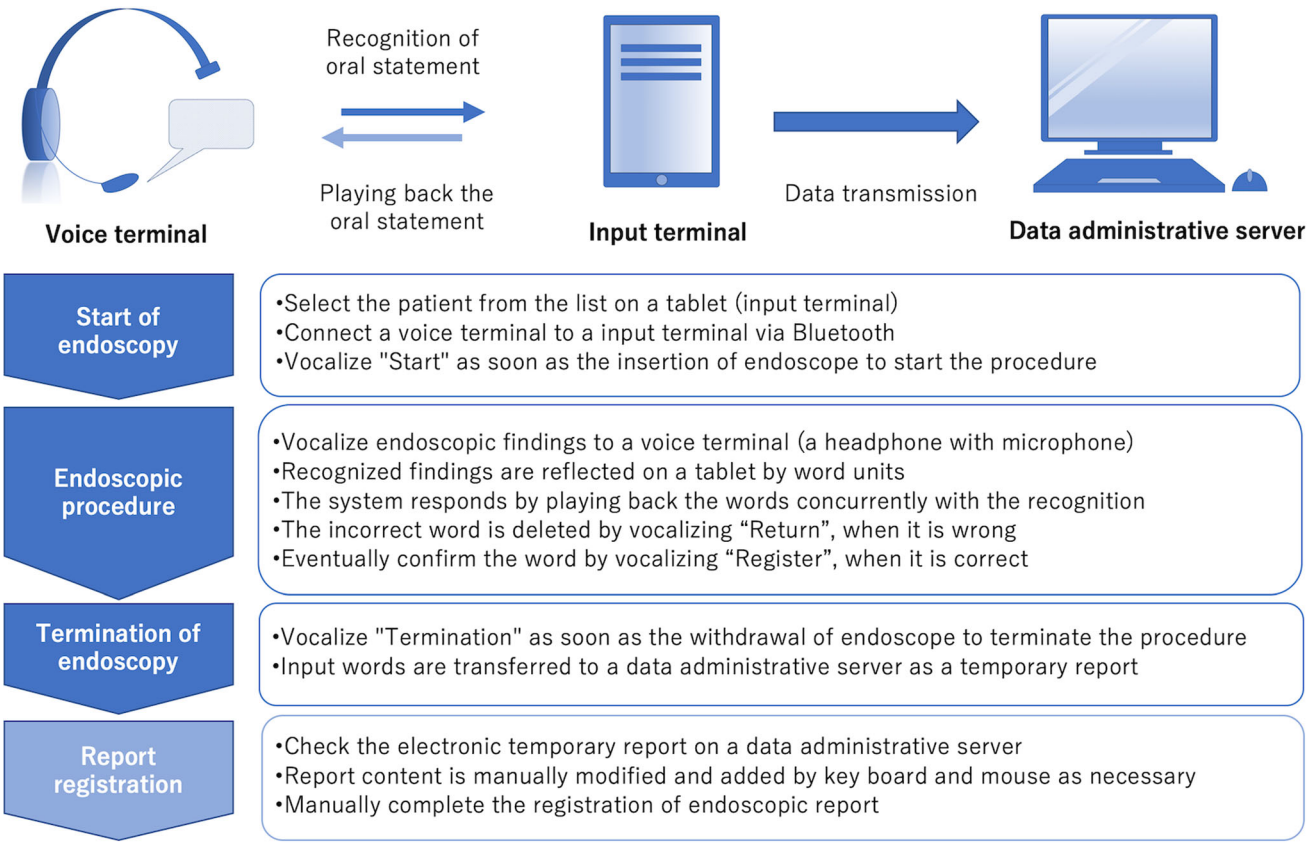

Fig. 1 A brief summary of report preparation using the endoscopic reporting system with voice recognition technology

In developing this system, we have introduced several novel technologies. When we first developed it, it was necessary to select and vocalize one word among the words shown on a display of a tablet. This meant that one needed to look at the tablet while performing the procedure. Thus, the new system was equipped with playing back the endoscopist's statement. This function allows the endoscopist to confirm whether the input words on the tablet are correct without looking at it (Fig. 2). Moreover, the voice recognition was not always accurate. The function of canceling surrounding noise was introduced to this system to recognize only the endoscopists' oral statements. This improved the speech recognition accuracy even with surrounding electric noise from biological data monitoring devices. The endoscopic findings can be selectively extracted from the recognized endoscopist's oral statements and automatically input, by word unit, to the appropriate place in the structured data. When the diagnosis name is unique to an organ, the organ name is simultaneously entered (e.g., if "gastric ulcer" is entered, it is automatically assigned to the "stomach" organ category). This function can reduce the endoscopist's vocalization frequency.

Outcomes

The primary endpoints were a comparison of the report preparation time and the total time spent on the procedure and report preparation between the VR and ME groups. The procedure time in the ME group was defined as the time from endoscope insertion to its withdrawal. The nurse or endoscopy technician manually recorded the time of endoscope insertion and withdrawal during the procedure using a mouse click and keyboard. In the VR group, the time of endoscope insertion and removal was recorded by the words "start" and "termination," respectively, vocalized by the endoscopist. The report preparation time in the two groups was measured according to the time recorded in the electronic information management system (from the receipt of transferred data to the report registration). The detected neoplastic lesions were classified based on endoscopic findings and histopathological diagnoses. When there were multiple neoplastic lesions in a patient, the most advanced lesion was adopted for analysis.

Statistical methods

We carried out propensity score matching to control and reduce selection bias in each case. Six variables that could possibly influence the total time spent on the endoscopic procedure and report preparation, including age, sex (male or female), purpose of EGD (screening, surveillance for gastrointestinal disease, or examination for gastrointestinal symptoms), history of endoscopic therapy (yes or no), under sedation (no, midazolam, or midazolam and pethidine), and endoscopists who performed the procedure (A, B, C, D, E, F, or G), were used to generate a propensity score by logistic regression [9–15]. We created a propensity score-matched cohort by attempting to match a patient who underwent EGD reported by VR with a patient who underwent EGD reported by ME (1:1 match), using a nearest neighbor matching without replacement. A caliper

Fig. 2 Endoscopic reporting system with voice recognition technology during the procedure. This image shows the voice recognition system used for report preparation during the procedure. An endoscopist wearing a headphone with a microphone is pronouncing the endoscopic finding while looking at an endoscopic picture on a monitor (red arrow). Although the words entered by oral statements (vocalization) are shown on a tablet (yellow arrow), the playback function allows the endoscopist to confirm whether the entered words are correct without looking at the tablet

width of 0.2 of the standard deviation (SD) of the logit of the propensity score was used for the developed propensity score. We used the standardized difference to measure covariate balance, whereby an absolute standardized difference (ASD) > 0.25 represents a meaningful imbalance [16].

Continuous variables are presented as the mean and SD or median and interquartile range, as appropriate for the data type. The Chi-squared or Fisher's exact tests were used to analyze categorical variables. Probability values for statistical tests were two-tailed, and a $P < 0.05$ was considered significant. Multiple linear regression analysis was performed to estimate the relationship between the reporting method (ME or VR) and the total, procedure, and report preparation times. The following factors that could be associated with report preparation or procedure times were included in the analyses: purpose of EGD (screening, surveillance for gastrointestinal disease, or examination for gastrointestinal symptoms); endoscopists (A, B, C, D, E, F, or G); under sedation (no, midazolam or midazolam and pethidine); chromoendoscopy (yes or no); the number of endoscopic findings; the number of biopsies; and the number of malignant lesion [10–15, 17, 18]. All analyses were performed using the statistical program "R," version 4.0.4 (R Foundation, Vienna, Austria).

Results

A total of 1492 consecutive patients underwent EGD from September 2019 to March 2020 (Fig. 3). We retrospectively reviewed 356 patients (130 patients in the VR group and 226 patients in the ME group) after exclusion of the rest from analysis for the following reasons: 880 patients underwent examinations by endoscopy systems from companies not supporting a VR system, 143 patients underwent EGD by endoscopists who had never used a VR system, 29 patients had a history of surgical treatment of the upper gastrointestinal tract (6 patients in the VR group and 23 patients in the ME group), 44 patients underwent EGD with an unknown examination time (1 patient in the VR group and 43 patients in the ME group), and 40 patients underwent an endoscopic procedure or had a report preparation lasting beyond 60 min (16 patients in the VR group and 24 patients in the ME group). None of the 356 patients underwent therapeutic endoscopy. After propensity score matching, 101 pairs of patients in the VR group and the ME group were selected.

The background characteristics of the patients and EGDs are summarized in Table 1. The change of the ASD showed that the balance between the two groups improved. Comparisons of the outcomes between the VR and ME groups after propensity score matching are shown in Table 2. Malignant lesions and early neoplasms were not significantly different between the two groups. The total time in the VR group tended to be shorter than that in the ME group (765 vs. 842 s, $P = 0.053$). The report preparation time in the VR group was significantly shorter than that in the ME group (311 vs. 383 s, $P = 0.009$), whereas the procedure time in the VR group was comparable to that in the ME group (429 vs. 457 s, $P = 0.59$).

The results of the multiple linear regression analysis in the propensity-matched cohort are shown in Table 3. Compared with ME, the VR system independently

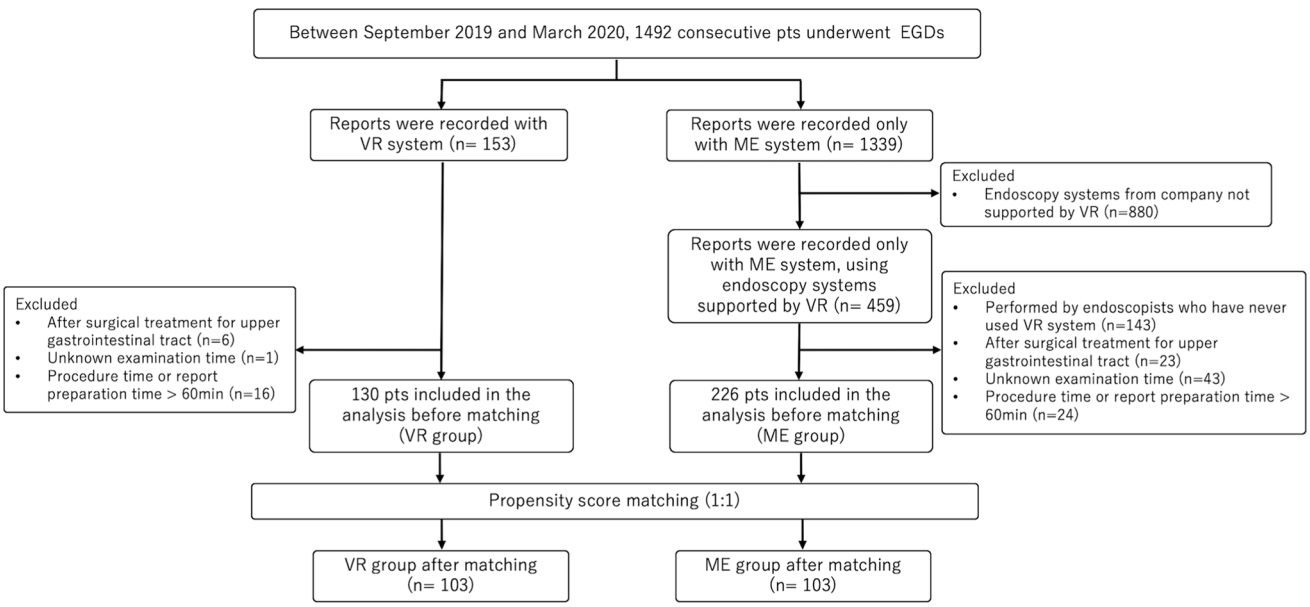

Fig. 3 Patient enrollment flow diagram. *EGD* esophagogastroduodenoscopy, *ME* manual entry, *Pts* patients, *VR* voice recognition

Table 1 Baseline characteristics of unmatched and propensity score-matched cases in the manual entry and voice recognition groups

	Unmatched			Propensity score-matched		
	ME (n = 226)	VR (n = 130)	ASD	ME (n = 101)	VR (n = 101)	ASD
Age, years, median [IQR]	71 [58.3–79.0]	61.5 [50.0–74.0]	0.44	65.0 [52.0–76.0]	67.0 [54.0–76.0]	0.06
Sex						
Male, n (%)	103 (45.6)	60 (46.2)	0.01	55 (54.5)	49 (48.5)	0.11
Female, n (%)	123 (54.4)	70 (53.8)		46 (45.5)	52 (51.5)	
Purpose of EGD						
Screening, n (%)	60 (26.5)	71 (54.6)	0.15	44 (43.6)	42 (41.6)	0.05
Surveillance for GI disease, n (%)	71 (31.4)	34 (26.2)		31 (30.7)	34 (33.7)	
Examination for GI symptoms, n (%)	95 (42.0)	25 (19.2)		26 (25.7)	25 (24.8)	
History of endoscopic therapy						
No, n (%)	214 (94.7)	127 (97.7)	0.64	97 (96.0)	98 (97.0)	0.06
Yes, n (%)	12 (5.3)	3(2.3)		4 (4.0)	3 (3.0)	
Esophageal EMR/ESD, n	1	3		0	3	
Gastric EMR/ESD, n	11	0		4	0	
Sedation						
No, n (%)	17 (7.5)	22 (16.9)	0.41	15 (14.9)	12 (11.9)	0.09
Yes, n (%)	209 (92.5)	107 (83.1)		88 (85.1)	89 (88.1)	
Midazolam, n (%)	156 (69.0)	65 (50.0)		60 (59.4)	61 (60.4)	
Midazolam and pethidine, n (%)	53 (23.5)	43 (33.1)		26 (25.7)	28 (27.7)	
Endoscopist						
A, n (%)	36 (15.9)	27 (20.8)	0.54	17 (16.8)	20 (19.8)	0.15
B, n (%)	36 (15.9)	43 (33.1)		22 (21.8)	23 (22.8)	
C, n (%)	27 (11.9)	16 (12.3)		16 (15.8)	15 (14.9)	
D, n (%)	20 (8.8)	5 (3.8)		7 (6.9)	5 (5.0)	
E, n (%)	37 (16.4)	14 (10.8)		17 (16.8)	14 (13.9)	
F, n (%)	14 (6.2)	9 (6.9)		6 (5.9)	8 (7.9)	
G, n (%)	56 (24.8)	16 (12.3)		16 (15.8)	16 (15.8)	

ASD absolute standardized difference, *IQR* interquartile range, *EGD* esophagogastroduodenoscopy, *GI* gastrointestinal, *EMR* endoscopic mucosal resection, *ESD* endoscopic submucosal dissection, *SD* standard deviation, *ME* manual entry, *VR* voice recognition

shortened the total time by 156 s (95% confidence interval [CI], − 274 to − 37 s; P = 0.009). Moreover, it independently shortened the report preparation time by 118 s (95% CI, − 220 to − 15 s; P = 0.023). However, the VR system was not significantly associated with the procedure time (estimate = − 38 s; 95% CI, − 77 to 1 s; P = 0.06).

Discussion

This study demonstrated the efficacy of VR technology to reduce the examination time in clinical endoscopy. The results of this study indicate that, compared with ME, the VR system reduced the total time spent on the procedure and report preparation. Furthermore, the VR system reduced the report preparation time, whereas the procedure time was not different between the two methods of report preparation. Thus, the reduction of report preparation time could have resulted in the reduction of the total time.

To date, a single clinical study on an endoscopic reporting system with a VR system has been reported; however, the reported system was used after the endoscopic procedure, and its results of report preparation time were not favorable [17]. Recently, a study evaluated VR using an EGD simulator [18]. To the best of our knowledge, the present study is the first report on a VR reporting system usable during clinical endoscopic procedures. This new system adopted the function of playing back the endoscopist's statements, unlike conventional systems, which means that the operator no longer needs to look at another screen, such as that of a tablet, to confirm the entered words. Therefore, this might prevent missing new

Table 2 Outcomes of esophagogastroduodenoscopy in the manual entry and voice recognition groups after propensity score matching

	ME (n = 101)	VR (n = 101)	P value
Chromoendoscopy			
No, n (%)	60 (59.4)	62 (61.4)	0.77
Yes, n (%)	41 (40.6)	39 (38.6)	
Indigo carmine, n	41	39	
Lugol, n	0	0	
No. of endoscopic findings per patient, median [IQR]	3 [2–3]	3 [2–4]	0.28
None, n (%)	3 (3.0)	2 (2.0)	
1–2, n (%)	43 (42.6)	42 (41.6)	
3–4, n (%)	50 (50.0)	42 (41.6)	
$\geq$ 5, n (%)	5 (5.0)	15 (14.9)	
No. of biopsies per patient, median [IQR]	0 [0–1]	0 [0–1]	0.84
None, n (%)	64 (63.3)	64 (63.3)	
1, n (%)	26 (25.7)	29 (28.7)	
2, n (%)	6 (5.9)	7 (6.9)	
$\geq$ 3, n (%)	5 (5.0)	1 (1.0)	
Patients with malignant lesions, n (%)	3 (3.0)	3 (3.0)	1.00
Esophageal neoplasms, n	1	0	
Gastric neoplasms, n	2	2	
Duodenal neoplasms, n	0	0	
Patients with early neoplasm, n (%)	1 (1.0)	2 (2.0)	1.00
Superficial esophageal cancer, n	0	0	
Gastric adenoma, n	1	1	
Early gastric cancer, n	0	1	
Total time, median [IQR], s	842 [634, 1286]	765 [607, 1064]	0.053
Procedure time, median [IQR], s	457 [326, 632]	429 [336, 577]	0.59
Reporting time, median [IQR], s	383 [263, 587]	311 [228, 412]	0.009

IQR interquartile range, *EGD* esophagogastroduodenoscopy, *ME* manual entry, *VR* voice recognition

lesions during the procedure because the endoscopists will be concentrating their eyes on the endoscopic monitor. This study found no differences in the number of detected malignant lesions, especially early-stage neoplasms, between the two groups. Thus, this system is unlikely to cause a higher occurrence of missed lesions during the procedure, although further investigation is required because of the small sample size of the present study.

The multivariate analysis indicated that the VR system reduced the report preparation time but barely affected the procedure time. There was a possibility of time extension due to repetition of oral statements if the recognition accuracy was low. However, there was no tendency to prolong the procedure time. A questionnaire survey showed that most EGD procedures lasted 5–8 min in Japanese hospitals [19]. Furthermore, Kawamura et al. reported that the average procedure time for screening EGDs was 6.2 min [12]. The procedure time in the current study was consistent with these previous findings. Furthermore, it was considered reasonable, considering that more than half of the EGDs were performed for surveillance for gastrointestinal diseases or examination for gastrointestinal symptoms. Therefore, the VR system does not seem to prolong the endoscopic procedure time.

The VR technology is now being applied to medical documentation, such as electronic health records and pathology and radiology reports, and has been found to decrease recording and report preparation times [5, 6, 20, 21]. However, the most crucial difference between endoscopy reports and other medical tools lies in the ability to create reports while performing the examination. Conventional endoscopy reports using ME must be typed and written after the procedure. Therefore, compared with other medical tools, the ability of the VR system to create reports in parallel with the examination can reduce the report preparation time. In this study, the use of the VR system shortened the preparation time and the total time by 117 and 156 s, respectively. This means that approximately one-third of the report preparation time and one-fifth of the total time were saved. In a situation where the endoscopist has to perform a large number of endoscopies on a daily basis, prolonged examination times can lead to fatigue and

Table 3 The effect of factors on total time, procedure time, and report preparation time: multiple linear regression analysis in the propensity-matched cohort

	Total time		Procedure time		Report preparation time	
	Estimate [95%CI], s	*P* value	Estimate [95%CI], s	*P* value	Estimate [95%CI], s	*P* value
Reporting method						
ME	Reference		Reference		Reference	
VR	− 156 [− 274, − 37]	0.009	− 38 [− 77, 1]	0.060	− 118 [− 220, − 15]	0.023
Sedation						
No	Reference		Reference		Reference	
Midazolam	− 84 [− 268, 99]	0.36	0.4 [− 32, 114]	0.99	− 85 [− 245, 74]	0.29
Midazolam and pethidine	− 78 [− 298, 140]	0.48	40 [− 32, 114]	0.27	− 119 [− 310, 71]	0.22
Chromoendoscopy						
No	Reference		Reference		Reference	
Yes	− 35 [− 99, − 170]	0.61	129 [84, 174]	< 0.001	− 94 [− 211, 22]	0.11
Purpose of EGD						
Screening	Reference		Reference		Reference	
Surveillance for GI disease	− 37 [− 189, 114]	0.63	26 [− 24, 77]	0.31	− 63 [− 195, 68]	0.34
Examination for GI symptoms	1 [− 156, − 159]	0.99	68 [15, 121]	0.010	− 66 [− 204, 70]	0.34
No. of biopsies per patient	112 [22, 203]	0.010	78 [48, 108]	< 0.001	34 [− 44, 113]	0.39
No. of endoscopic findings per patient	61 [5, 118]	0.031	11 [− 7, 30]	0.24	50 [1, 100]	0.042
No. of malignant lesions	88 [− 257, 452]	0.63	210 [89, 331]	< 0.001	− 122 [− 438, 193]	0.45

The reporting method (manual entry or voice recognition), purpose of EGD (screening, surveillance for gastrointestinal disease or examination for gastrointestinal symptoms), sedation (no, midazolam, or midazolam and pethidine), endoscopist (seven different operators), chromoendoscopy (yes or no), the number of endoscopic findings, the number of biopsies, and the number of malignant lesion are adjusted in the analysis

CI confidence interval, *EGD* esophagogastroduodenoscopy, *GI* gastrointestinal, *ME* manual entry, *VR* voice recognition

mental stress. Such a burden can also lead to a decrease in the quality of the examination. A previous study indicated that the increase in the number of screening examinations might increase endoscopists' workloads and consequently reduce gastric cancer detection rates [22]. The adenoma detection rate (ADR) in colonoscopy tends to decrease later in the day (increasing procedure queue position) [23, 24]. Several reports have suggested that endoscopist fatigue may reduce the ADR [23, 25, 26]; therefore, the VR system may mitigate the burden of examination work on the endoscopist. To expand the application of this system beyond EGD, we are currently developing the same system for colonoscopy, and an English version is under consideration. This reporting system can be upgraded and customized in multiple ways, in accordance with endoscopists' needs. Further studies and developments of endoscopic reporting systems with VR are warranted in the future to assist in endoscopy-related tasks.

This study had several limitations. First, this was a single-center, retrospective study. Although the endoscopists were not informed of the aim of this study during the study, they could choose either method of report preparation depending on the patient's background, such as the purpose of the examination. Therefore, selection bias was potentially present. We performed propensity score matching to address this limitation. Second, this study focused on routine EGD other than complex procedures, such as endoscopic treatment. The applicability of the results of this study to endoscopic procedures or colonoscopy other than routine EGD is unknown. Third, the learning curve of this system was not investigated. The number of cases required to reach a plateau in the examination and report times using this system is unknown. Therefore, if some of the cases included in this study fell on the learning curve, the results might have been an underestimation.

In conclusion, our novel reporting system using VR reduced the report preparation time and the total time spent on endoscopic procedures and report preparation. This novel system might improve the efficiency of report preparation and enable endoscopists to spend more time on other beneficial tasks.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00535-021-01835-7.

Acknowledgements The authors thank the medical staff of the endoscopy unit of the Hanwa Sumiyoshi General Hospital. We would

like to thank Editage (www.editage.com) for English language editing. Moreover, we would like to express our sincere gratitude to Prof. Ervin Toth (Department of Gastroenterology, Skåne University Hospital, Lund University, Malmö, Sweden) for his valuable comments and suggestions in preparing this paper. There was no funding support for this study.

Author contributions YY, TI, and TT, conception and design. YY, AH, YN, KK, YI, HM, KI, JM, and TA, investigation. TI and AH, formal analysis. YY and TI, drafting of the article. TI and TT, critical revision of the article for important intellectual content. All authors, final approval of the article.

Declarations

Conflict of interest Authors declare no conflict of interests for this article.

References

1. Bretthauer M, Aabakken L, Dekker E, et al. Requirements and standards facilitating quality improvement for reporting systems in gastrointestinal endoscopy: European society of gastrointestinal endoscopy (ESGE) position statement. Endoscopy. 2016;48:291–4.
2. Delvaux M, Korman LY, Armengol-Miro JR, et al. The minimal standard terminology for digestive endoscopy: introduction to structured reporting. Int J Med Inform. 1998;48:217–25.
3. Gouveia-Oliveira A, Raposo VD, Salgado NC, et al. Longitudinal comparative study on the influence of computers on reporting of clinical data. Endoscopy. 1991;23:334–7.
4. Manfredi MA, Chauhan SS, ASGE Technology Committee, et al. Endoscopic electronic medical record systems. Gastrointest Endosc. 2016;83:29–36.
5. Krishnaraj A, Lee JKT, Laws SA, et al. Voice recognition software: effect on radiology report turnaround time at an academic medical center. Am J Roentgenol. 2010;195:194–7.
6. Shimazui T, Nakada TA, Kuroiwa S, et al. Speech recognition shortens the recording time of prehospital medical documentation. Am J Emerg Med. 2021. https://doi.org/10.1016/j.ajem.2021.02.025.
7. Matsuda K, Tanaka K, Fujishiro M, et al. Design paper: Japan endoscopy database (JED): a prospective, large database project related to gastroenterological endoscopy in Japan. Dig Endosc. 2018;30:5–19.
8. Japan Gastroenterological Endoscopy Society. The Japan Endoscopy Database (JED) project screening JED upper gastrointestinal terminology version 1.2. Available from: https://jedproject.jges.net/about/terms-download. Accessed 31 May 2021 (**in Japanese**)
9. Mori G, Nakajima T, Asada K, et al. Incidence of and risk factors for metachronous gastric cancer after endoscopic resection and successful *Helicobacter pylori* eradication: results of a large-scale, multicenter cohort study in Japan. Gastric Cancer. 2016;19:911–8.
10. Katada C, Horimatsu T, Muto M, et al. Current status of esophageal endoscopy including the evaluation of smoking and alcohol consumption in Japan: an analysis based on the Japan endoscopy database. Esophagus. 2019;16:174–9.
11. Gotoda T, Akamatsu T, Abe S, et al. Guidelines for sedation in gastroenterological endoscopy (second edition). Dig Endosc. 2021;33:21–53.
12. Kawamura T, Wada H, Sakiyama N, et al. Examination time as a quality indicator of screening upper gastrointestinal endoscopy for asymptomatic examinees. Dig Endosc. 2017;29:569–75.
13. Park JM, Huo SM, Lee HH, et al. Longer observation time increases proportion of neoplasms detected by esophagogastroduodenoscopy. Gastroenterology. 2017;153:460-469.e1.
14. Teh JL, Tan JR, Lau LJF, et al. Longer examination time improves detection of gastric cancer during diagnostic upper gastrointestinal endoscopy. Clin Gastroenterol Hepatol. 2015;13:480-487.e2.
15. Park JM, Kim SY, Shin GY, et al. Implementation effect of institutional policy of EGD observation time on neoplasm detection. Gastrointest Endosc. 2021;93:1152–9.
16. Stuart EA. Matching methods for causal inference: a review and a look forward. Stat Sci. 2010;25:1–21.
17. Massey BT, Geenen JE, Hogan WJ. Evaluation of a voice recognition system for generation of therapeutic ERCP reports. Gastrointest Endosc. 1991;37:617–20.
18. Takao T, Masumura R, Sakauchi S, et al. New report preparation system for endoscopic procedures using speech recognition technology. Endosc Int Open. 2018;6:E676–87.
19. Uedo N, Gotoda T, Yoshinaga S, et al. Differences in routine esophagogastroduodenoscopy between Japanese and international facilities: a questionnaire survey. Dig Endosc. 2016;28(Suppl 1):16–24.
20. Kang HP, Sirintrapun SJ, Nestler RJ, et al. Experience with voice recognition in surgical pathology at a large academic multi-institutional center. Am J Clin Pathol. 2010;133:156–9.
21. Vogel M, Kaisers W, Wassmuth R, et al. Analysis of documentation speed using web-based medical speech recognition technology: randomized controlled trial. J Med Internet Res. 2015;17:e247.
22. Noh CK, Lee E, Lee GH, et al. Association of intensive endoscopic screening burden with gastric cancer detection. JAMA Netw Open. 2021;4:e2032542.
23. Lee TJW, Rees CJ, Blanks RG, et al. Colonoscopic factors associated with adenoma detection in a national colorectal cancer screening program. Endoscopy. 2014;46:203–11.
24. Marcondes FO, Gourevitch RA, Schoen RE, et al. Adenoma detection rate falls at the end of the day in a large multi-site sample. Dig Dis Sci. 2018;63:856–9.
25. Lee A, Iskander JM, Gupta N, et al. Queue position in the endoscopic schedule impacts effectiveness of colonoscopy. Am J Gastroenterol. 2011;106:1457–65.
26. Lee CK, Cha JM, Kim WJ. Endoscopist fatigue may contribute to a decline in the effectiveness of screening colonoscopy. J Clin Gastroenterol. 2015;49:e51–6.